A Time-Varying Structural VAR Model to Estimate the Effects of Changes in Fiscal Policy

By Klemens Hauzenberger*

November 11, 2011

In this paper I study the dynamic macroeconomic effects of changes in government spending or taxes by means of a time-varying structural VAR. Such a model was used by Giorgio E. Primiceri (2005) to analyze monetary policy but is, with a few exceptions, novel to address the question of the transmission mechanism of fiscal policy. My results accord well with the notion of important changes in that mechanism over time: the effectiveness of fiscal policy in stabilizing the economy has decreased, more or less so for tax shocks and de facto with respect to spending. I also find evidence, through counterfactual policy simulations, for positive longer run effects on output when the government actively reduces the level of debts by cutting spending. A passive reduction of debts through faster tax adjustments in response to past expenditures has adverse effects on output. (JEL E62; H30; H50; C32; C53)

Keywords: fiscal policy; transmission of shocks; policy counterfactuals; time-varying parameter structural VAR; Markov chain Monte Carlo

1. Introduction

What are the effects of fiscal policy? The longest and the deepest recession since the Great Depression that started in December 2007 and the various stimulus and reinvestment measures enacted by the U.S. Congress to facilitate the recovery have ended the eclipse of fiscal policy

* Hauzenberger: Deutsche Bundesbank (Macroeconomic Analysis and Projection Division), Wilhelm-Epstein-Strasse 14, 60431 Frankfurt/Main (e-mail: klemens.hauzenberger@bundesbank.de). I gratefully acknowledge the advice and comments of Helmut Lütkepohl, Massimiliano Marcellino and Simon van Norden. The views expressed in this paper are solely mine and should not be interpreted as reflecting the views of the Deutsche Bundesbank.

in the economic literature. Many economists from various fields of the profession have returned their interest to this classical question.

During my own research on the empirical effects of fiscal policy shocks over the last years, I have encountered two main road blocks on the way, which make precise and unbiased inference difficult: the identification of shocks, especially the tax shocks,[1] and the state-dependent or changing transmission mechanism.[2] These issues leave, of course, plenty of room for new econometric approaches. To be clear where this paper is heading, the objective is to provide one possible solution for one of these road blocks: the changing transmission mechanism and the resulting differences in the effectiveness of fiscal policy. The particular application is to the U.S. economy over the 1970:1-2010:3 period.

The proposed solution is the time-varying parameter structural vector autoregressive (TVP-VAR) model of Primiceri (2005) in which the vector of variables includes measures for government spending, tax revenues, output, and federal debts. In the TVP-VAR model all coefficients, covariances and volatilities vary over time. The laws of motion for the parameters allow for smooth changes and, as such, the method differs from Markov switching or threshold models. These models are better suited to study fiscal policy in recessions and expansions rather than the evolution of its effectiveness. The estimation of the TVP-VAR is in the Bayesian tradition of Markov chain Monte Carlo algorithms, Gibbs sampling in particular, for the numerical evaluation of the posterior distributions of the parameters.

To identify government spending and tax shocks I follow Olivier Blanchard and Perotti (2002). Their method is the most compelling and cited VAR-based approach to identifying fiscal policy shocks and I will take it as given throughout this paper. Simply put, identification rests on the ability to disentangle, from the residual changes in government spending and revenues, the discretionary part (i.e. the shocks) and the automatic adjustment to output. The difficulty, then, and methodological innovation of this paper is to cast this essentially non-recursive identification strategy into the one of Primiceri (2005), a strategy that relies on a triangular structure of the identifying matrix in order to preserve the assumption of a linear state space model for the TVP-VAR.

[1] See among others the seminal papers of Christina D. Romer and David H. Romer (2010) and Robert Perotti (2011) for measures of tax changes based on the narrative record of all major postwar tax policy acts; and Eric M. Leeper, Todd B. Walker and Shu-Chun Susan Yang (2008) on the problems arising through fiscal foresight.

[2] Papers in this direction are Alan Auerbach and Yuriy Gorodnichenko (2010), Markus Kirchner, Jacopo Cimadomo and Sebastian Hauptmeier (2010) and Manuel Coutinho Pereira and Artur Silva Lopes (2010).

The paper further innovates in a second direction. Besides tracing out the effects of the policy shocks, I estimate impulse response functions of changes in parameters of the government's decision rules on spending and taxes. Specifically, I simulate a government that tries to reduce the level of debts in two different ways: in an active one by cutting spending and in a passive one by just adjusting taxes to cover past expenditures. The design of the policy counterfactuals borrows from the monetary policy literature, especially from work of Fabio Canova and Luca Gambetti (2009). Their approach is particularly appealing because it takes the Lucas critique seriously by taking the estimated covariance structure of the coefficients into account in the experiment. Moreover, including the federal debts in the vector of variables, controls for the constraints the debt path puts on future spending and tax decisions. A channel typically ignored in the VAR-based fiscal policy analysis (see Carlo Favero and Francesco Giavazzi, 2007).

While there is a huge recent literature using TVP-VARs to evaluate monetary policy (see, e.g., Primiceri, 2005; Timothy Cogley and Thomas J. Sargent, 2001; Canova and Gambetti, 2009; Luca Benati and Paolo Surico, 2008), applications for fiscal policy are scarce. Two notable exceptions are Kirchner, Cimadomo and Hauptmeier (2010) and Pereira and Lopes (2010). The first of these two papers traces out the effects of government spending shocks in the euro area and the second one identifies both spending and tax shocks for the postwar U.S. economy. Both papers confirm the notion of important changes in the transmission mechanism over time: the effectiveness of fiscal policy in stabilizing the economy has decreased on both sides of the Atlantic.

The first set of my results is much in line with this general finding of Kirchner, Cimadomo and Hauptmeier (2010) and Pereira and Lopes (2010). Changes in government spending had a stronger positive effect on output, especially after six or seven quarters, in the 1970s and early 1980s. The picture is a bit different for the responses to tax shocks. Still, one can observe the same pattern between the 1970s and 2000s, but the late 1980s and early 1990s now seem to be the period when tax shocks were most effective. Unlike the other periods, this mid-period was mainly characteristic for a few deficit-driven shocks (see Romer and Romer, 2010), with a more persistent response of tax revenues and the desired effect of a significant reduction of federal debts.

The results from the debt-reducing counterfactual policies suggest that the spending cuts in the active government stance have hardly any adverse effects on the private sector and output increases over time. Not surprisingly, just levying taxes to achieve budget surpluses without changing the spending behavior has detrimental effects on output.

2. Econometric Framework

Time-varying parameter structural VARs put quite a challenge on an econometrician because of the sheer amount of parameters to estimate: 8,732(!) to be exact in the model presented in this section. While it is still possible to write down the likelihood for the estimation problem, it comes close to a mission impossible to maximize it over such a high dimension, let alone the problem of multiple maxima in ranges where the parameter values are anything but plausible. This classical approach to estimation is basically a special case of a Bayesian one with flat priors. Bayesian estimation with informative or diffuse priors is therefore the natural choice to tackle the problem. Section 3 has the details.

Compared to the unproblematic and relative uncontroversial identification of monetary policy, disentangling fiscal policy shocks is far from trivial. Because, strictly speaking, there is no such thing as a “universal” fiscal shock that accounts for the numerous strings policy makers can pull to counteract the business cycle by changes in spending and taxes. A billion dollars spent for public infrastructure, education, or defense will hardly have the same effects both on the individual citizen or the economy as a whole. In this paper I am, however, pragmatic about this problem and keep the focus on the traditional macroeconomic issue of the aggregate economy. Focusing on the aggregate economy and, by implication, on *total* government spending and tax revenue shocks is in line with seminal papers such as Blanchard and Perotti (2002) and Andrew Mountford and Harald Uhlig (2009).

Even though the traditional macroeconomic approach simplifies matters considerably, identifying fiscal policy shocks remains difficult because of endogeneities. Fiscal variables and the business cycle are closely linked. For instance, under a fixed tax code both higher taxes and higher economic growth will fill the Federal Treasury and, as a consequence, we do not know whether the rise in tax revenues comes from a tax or business cycle shock. To overcome this difficulty I use the method of Blanchard and Perotti (2002) to identify spending and tax shocks. At the heart of their structural VAR methodology lies the identification of the just sketched automatic “feedback” of economic activity on tax revenues and government spending. The identification rests on additional information from outside the VAR model about the tax and transfer system in order to pin down these feedback elasticities. It also relies on a timing assumption. The quasi-impossibility of any discretionary within-quarter adjustment of fiscal policy in response to economic shocks attributes any contemporaneous changes to the feedback effects.

2.1. Data Description

The sample covers quarterly observations for the United States from 1970:1 until 2010:3. To keep my results, with respect to the definition of the fiscal data, comparable with the seminal work of Blanchard and Perotti (2002) and other studies in their tradition, I define these variables accordingly: spending includes both government consumption expenditures and gross investment, and net taxes are the current receipts less net transfers and net interest paid. The model further includes data on output and federal debts. As Favero and Giavazzi (2007) forcefully argue and show, it is important to control for the restrictions the debt path puts on future government spending and tax decisions. Ignoring this channel would potentially bias the estimates of the fiscal policy effects.

All variables enter the analysis in logarithmic form of their respective real per capita values. The sources for nominal output, government spending, the necessary items to construct net taxes, the deflator, population and federal debts are the NIPA tables and the FRED data base.[3]

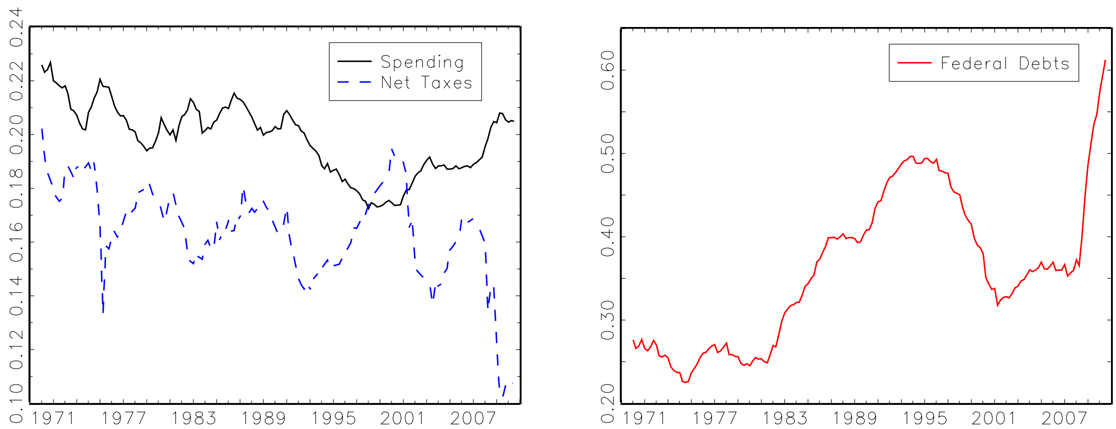

FIGURE 1. SPENDING, NET TAX AND DEBT SHARES OF OUTPUT

Figure 1 shows some of the high frequency properties of the data. While, in the econometric analysis, I use the variables in the level-form described above, for a quick visual inspection it is more appealing to look at, say, the shares of spending, net taxes and federal debts with respect to output. Most importantly, the shares display patterns that make the time-varying parameter model a natural choice. For instance, the spending share declines steadily until the beginning

[3] Specifically, NIPA tables 1.1.4 (line 1), 1.1.5 (lines 1 and 21), 3.1 (lines 1, 9, 11, 17 and 22), 7.1 (line 18); and FRED data base (series FYGFPUN). All variables downloaded on February 21, 2011.

of the new millennium but increase since then. Likewise, the debt share remains either stable, increases, or decreases over periods of several years. Besides these “long swings”, which can be perfectly captured by TVP-VARs, there are a couple of large quarterly changes in taxes and federal debts. The first episode is President Ford’s large temporary tax rebate of 1975:2 and the second one are the recent effects of the Great Recession. TVP-VARs typically detect such episodes of large and quick changes only rudimentary and capturing such changes is not the objective of this paper. Intuitively, the more time variation one allows for in the VAR the more will be explained by the shocks as opposed to the dynamics of the model.

2.2. Model Specification

The k-dimensional vector of quarterly observable variables, $\{y_t\}_{t=1}^{T}$, includes government spending, net taxes, output, and federal debts in this order. I assume $y_t = (y_{g,t},\, y_{t,t},\, y_{x,t},\, y_{d,t})'$ evolves according to the TVP-VAR(p) process,

$$y_t = C_t + B_{1,t}y_{t-1} + \cdots + B_{p,t}y_{t-p} + u_t, \tag{1}$$

in which C_t is a $k \times 1$ vector of time-varying intercepts, $B_{i,t}$ $(i = 1, \ldots, p)$ are $k \times k$ matrices of time-varying coefficients and u_t are possibly heteroscedastic reduced-form residuals with time-varying covariance matrix Ω_t. Iterating on (1) yields the corresponding infinite moving average representation, i.e.

$$y_t = \mu_t + \sum_{h=1}^{\infty} \Theta_{h,t} u_{t-h}. \tag{2}$$

$\mu_t = I_k + \sum_{h=1}^{\infty} \Theta_{h,t} C_t$ and $\Theta_{h,t} = J\tilde{B}_t^h J'$ in which $\tilde{B}_t$ is the corresponding VAR(1) companion form of the VAR(p) in (1) and J a selector matrix:

$$\tilde{B}_t = \begin{bmatrix} B_t \\ I_{k(p-1)} : 0_{k(p-1)\times k} \end{bmatrix} \quad \text{and } J = \left(I_k : 0_{k\times k(p-1)}\right). \tag{3}$$

The parameters $\Theta_{h,t}$ for $h = 1, \ldots, H$ represent the reduced-form impulse response functions. To transform these responses into ones with a structural interpretation I use the fairly general model of Blanchard and Perotti (2002) that links the reduced-form residuals u_t with the structural shocks e_t:

$$(4)\qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\alpha^*_{23,t} & 0 \\ -\alpha_{31,t} & -\alpha_{32,t} & 1 & 0 \\ -\alpha_{41,t} & -\alpha_{42,t} & -\alpha_{43,t} & 1 \end{bmatrix} \begin{bmatrix} u_{g,t} \\ u_{t,t} \\ u_{x,t} \\ u_{d,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \alpha_{21,t} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_{g,t} \\ e_{t,t} \\ e_{x,t} \\ e_{d,t} \end{bmatrix},$$

in which (minus) $\alpha^*_{23,t}$ is the predetermined tax elasticity with respect to output. Following Favero and Giavazzi (2007) and Perotti (2007) I use a value of 1.85 for this elasticity for all $t = 1, \ldots, T$ and, implicitly, a zero spending elasticity.[4] Blanchard and Perotti (2002) have the details on how to construct these elasticities based on institutional information outside the information set of the VAR.

The non-recursive structure of the Blanchard-Perotti model, however, imposes a twist on the time-varying parameter framework of Primiceri (2005). In his model, identification relies on a recursive Choleski-like decomposition and consequently on a lower triangular matrix linking the reduced-form residuals with the structural shocks. Now, the specific form of (4) allows me to recast the problem into a lower triangular matrix that contains all the parameters we want to estimate and a second matrix that collects the remaining predetermined variables.[5] Specifically,

$$(5)\qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\alpha^*_{23,t} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_{g,t} \\ u_{t,t} \\ u_{x,t} \\ u_{d,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \tilde{\alpha}_{21,t} & 1 & 0 & 0 \\ \tilde{\alpha}_{31,t} & \tilde{\alpha}_{32,t} & 1 & 0 \\ \tilde{\alpha}_{41,t} & \tilde{\alpha}_{42,t} & \tilde{\alpha}_{43,t} & 1 \end{bmatrix} \begin{bmatrix} e_{g,t} \\ e_{t,t} \\ e_{x,t} \\ e_{d,t} \end{bmatrix},$$

in which the mapping between (4) and (5) is

$$(6)\qquad \begin{aligned} \tilde{\alpha}_{21,t} &= \alpha_{21,t}, \\ \tilde{\alpha}_{31,t} &= \alpha_{31,t} + \alpha_{32,t}\alpha_{21,t}, \\ \tilde{\alpha}_{32,t} &= \alpha_{32,t}, \\ \tilde{\alpha}_{41,t} &= \alpha_{41,t} + \alpha_{42,t}\alpha_{21,t} + \alpha_{43,t}\alpha_{31,t} + \alpha_{43,t}\alpha_{32,t}\alpha_{21,t}, \\ \tilde{\alpha}_{42,t} &= \alpha_{42,t} + \alpha_{43,t}\alpha_{32,t}, \quad \text{and} \end{aligned}$$

[4] The fact that government spending, defined as in Section 2.1, does not include transfer payments justifies the assumption of no feedback effect of spending to movements in the business cycle. See also the evidence in Blanchard and Perotti (2002).

[5] This way to reparameterize (4) follows in principle the idea of Pereira and Lopes (2010).

$$\tilde{\alpha}_{43,t} = \alpha_{43,t}.$$

We can write the structural model more compactly as

$$A_t^* u_t = \tilde{A}_t \Sigma_t \varepsilon_t. \tag{7}$$

ε_t are the normalized structural shocks, i.e. $\text{Var}(e_t) = \Sigma_t$ is the diagonal matrix

$$\Sigma_t = \begin{bmatrix} \sigma_{1,t} & 0 & 0 & 0 \\ 0 & \sigma_{2,t} & 0 & 0 \\ 0 & 0 & \sigma_{3,t} & 0 \\ 0 & 0 & 0 & \sigma_{4,t} \end{bmatrix}, \tag{8}$$

which, in turn, leads to a reduction of the reduced-form covariance matrix Ω_t given by

$$\Omega_t = A_t^{*-1} \tilde{A}_t \Sigma_t \Sigma_t' \tilde{A}_t' A_t^{*-1\prime}. \tag{9}$$

The structural impulse responses follow then from

$$\Phi_{h,t} = \Theta_{h,t} A_t^{*-1} \tilde{A}_t \Sigma_t, \qquad h = 1, \ldots, H. \tag{10}$$

For the estimation it will be practical to collect the slope coefficients $B_t = (B_{1,t} : \cdots : B_{p,t})$ in a $k \times kp$ matrix and to transform it together with the constants into a $k(kp+1)$ vector by stacking the columns, i.e. $\beta_t = \text{vec}\left((C_t : B_t)'\right)$. The model (1) can now be rewritten as

$$\begin{aligned} y_t &= X_t' \beta_t + A_t^{*-1} \tilde{A}_t \Sigma_t \varepsilon_t, \\ X_t' &= I_k \otimes \left(1 : y_{t-1}' : \cdots : y_{t-p}'\right), \end{aligned} \tag{11}$$

in which the operator $\otimes$ denotes the Kronecker product. Like the constants and the slope coefficients, I bring the non-zero and non-one elements of the covariances $\tilde{A}_t$ and volatilities Σ_t into vector form. Specifically, $\tilde{\alpha}_t = (\tilde{\alpha}_{21,t}, \tilde{\alpha}_{31,t}, \tilde{\alpha}_{32,t} \cdots, \tilde{\alpha}_{k1,t}, \cdots, \alpha_{kk-1,t})'$ and $\sigma_t = (\sigma_{1,t}, \cdots, \sigma_{k,t})'$ where the corresponding dimensions are $k(k-1)/2 \times 1$ and $k \times 1$.

The vectors α_t, β_t, and σ_t summarize all the time-varying parameters of the model. As in Primiceri (2005) I let the coefficients α_t and β_t evolve as random walks and the standard deviations σ_t follow a geometric random walk:

$$\tilde{\alpha}_t = \tilde{\alpha}_{t-1} + \zeta_t \tag{12}$$

$$\beta_t = \beta_{t-1} + \nu_t, \tag{13}$$

$$\log \sigma_t = \log \sigma_{t-1} + \eta_t. \tag{14}$$

The specification for σ_t falls into the class of models known as stochastic volatility. While in infinite samples a random walk hits any bound for sure, the use of finite samples makes it possible to maintain the random walk assumption. A great advantage as we do not have to estimate any further parameters, although, in principle, we could extend (12), (13), and (14) to represent more general autoregressive processes (for details see Section 4.4.2 in Primiceri, 2005).

The innovations ε_t, ζ_t, ν_t, and η_t are mutually uncorrelated Gaussian white noises with zero mean and covariances I_k, Q, S, and W, known as the hyperparameters in the Bayesian literature. Summarized in the matrix V we have

$$V = \text{Var}\left(\begin{bmatrix} \varepsilon_t \\ \nu_t \\ \zeta_t \\ \eta_t \end{bmatrix}\right) = \begin{bmatrix} I_k & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix} \quad \text{and } S = \begin{bmatrix} S_1 & 0 & \cdots & 0 \\ 0 & S_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & S_{k-1} \end{bmatrix} \tag{15}$$

with all matrices, besides the k-dimensional identity matrix I_k, being positive definite. The covariance of the innovation term in the state equation for the log volatilities is block diagonal, i.e. $S_1 = \text{Var}([\Delta\tilde{\alpha}_{21,t}])$, and $S_{i-1} = \text{Var}([\Delta\tilde{\alpha}_{i1,t}, \cdots, \Delta\tilde{\alpha}_{ii-1,t}]')$ for all $i = 3, \ldots, k$. The rather specific assumptions on the structure of V and S are standard in the literature (see, e.g., Primiceri, 2005; Canova and Gambetti, 2009; Benati and Surico, 2008) and are not essential to keep the estimation feasible. They offer, however, numerous advantages: The first one is the clear structural interpretation of the various sources of uncertainty which would be cumbersome in the case of more non-zero blocks. Second, the block-diagonality of S with blocks corresponding to parameters in separate equations enables us to model the blocks $[\tilde{\alpha}_{21,t}]$, $[\tilde{\alpha}_{31,t}, \tilde{\alpha}_{32,t}], \ldots,$ $[\tilde{\alpha}_{k1,t}, \cdots, \tilde{\alpha}_{kk-1,t}]$ in linear state space form. The advantage of linearity will become clear momentarily in Section 3. And, finally, assuming mutually uncorrelated innovations does not exaggerated the curse-of-dimensionality problem inherent in all time-varying parameter models any further.

2.3. Counterfactual Fiscal Policy Scenarios

What would have happened to output had the government pursued a more aggressive policy to reduce to the level of debts? Questions like this appear to be simple with time series models: just change the relevant parameters in the decision rules and trace out the effects. But the effect of changing one parameter typically spreads over all the other parameters; the essence of the Lucas critique. Canova and Gambetti (2009) provide a natural solution to the critique, in an experiment that mimics a more aggressive monetary policy, by explicitly taking into account the covariance structure, i.e. the matrix Q in (15), of the whole coefficient set.

Let us define $G_t = A_t^{*-1}\tilde{A}_t\Sigma_t$ in (11) and rewrite the reduced-form model (1) in structural form,

$$G_t^{-1}y_t = G_t^{-1}C_t + G_t^{-1}B_{1,t}y_{t-1} + \cdots + G_t^{-1}B_{p,t}y_{t-p} + \varepsilon_t, \tag{16}$$

or equivalently

$$G_t^{-1}y_t = X_t'\left(G_t^{-1} \otimes I_{kp+1}\right)\beta_t + \varepsilon_t = X_t'\gamma_t + \varepsilon_t, \tag{17}$$

in which X_t' is defined as in (11) and γ_t are the structural coefficients. Using (13) and after some rearranging we get

$$\gamma_t = \left(G_t^{-1} \otimes I_{kp+1}\right)\left(G_{t-1}^{-1} \otimes I_{kp+1}\right)^{-1}\gamma_{t-1} + \left(G_t^{-1} \otimes I_{kp+1}\right)\nu_t \tag{18}$$

as the law of motion for the structural coefficients on the lagged variables. The last term, $\omega_t = \left(G_t^{-1} \otimes I_{kp+1}\right)\nu_t$, contains all the $k(kp+1)$ shocks. Let $\tilde{\omega}_t \subset \omega_t$ be the subvector containing the n shocks of interest and the submatrix $\tilde{G}_t$ consists of the n corresponding rows of $G_t^{-1} \otimes I_{kp+1}$ such that $\tilde{\omega}_t = \tilde{G}_t\nu_t$. Given the last expression and the covariance matrix Q from (13), we can write ν_t conditional on $\tilde{\omega}_t$ in turn as

$$\nu_t = Q\tilde{G}_t'\left(\tilde{G}_tQ\tilde{G}_t'\right)^{-1}\tilde{\omega}_t. \tag{19}$$

Now to compute the impulse response function of the policy counterfactual I use the method of Gary Koop, M. Hashem Pesaran and Simon M. Potter (1996) and take the difference between two different realizations of the forecast $E_t(y_{t+i}|\cdot)$. The two realizations are identical up to $t-1$, but one realization assumes that there is an innovation at date t of size $\tilde{\omega}_t = \delta$, while the other realization evolves along its regular innovation-free path. Specifically,

$$(20) \quad \Phi^c_{h,t} = E_t\left(y_{t+h} | \tilde{\omega}_t = \delta, \{\tilde{\omega}_{t+j}\}^h_{j=1} = 0, Q, \tilde{G}_t, \beta_{t-1}, \{y_{t-j}\}^p_{j=1}\right) - E_t\left(y_{t+i} | \{\tilde{\omega}_{t+j}\}^h_{j=0} = 0, Q, \tilde{G}_t, \beta_{t-1}, \{y_{t-j}\}^p_{j=1}\right)$$

for $h = 0, 1, 2, \ldots, H$. Although δ is only a one-time shock its effects are permanent through the random walk nature of the law of motion for the time-varying parameters. To further ensure against the Lucas critique, I calibrate δ to represent, in the sense of Leeper and Tao Zha (2003), only a modest or typical policy intervention in the sample. One posterior standard deviation of the corresponding shock ν_t to the reduced-form parameters β_t (i.e. the square root of the associated diagonal element of Q) is consistent with such a typical intervention.

My interest centers around two counterfactual fiscal policies that involve the structural equations for government spending and taxes. Similar to Taylor rules, these structural equations provide simple descriptions of fiscal policy-making, approximating the many complex mechanisms and constraints that influence the government's decisions. Now, in both counterfactual experiments, the objective is to bring the level of federal debts down. The ways to get there differ, however. In the first experiment, I simulate a government that actively pursues its objective by cutting spending more aggressively with respect to past debt levels. The other experiment shows what I call a passive government. It achieves the debt reduction not by actively reacting to debt levels as before. Rather, the government runs surpluses by adjusting taxes faster to recent expenditures.

For the technical implementation of the first experiment—the active government stance—I set up the matrix $\tilde{G}_t$ to contain the rows corresponding to the lagged coefficients of spending and debts in the spending equation. The specific shocks in the vector δ hit the debt coefficients by minus one-standard deviation and leave the spending coefficients (i.e. the autoregressive component) unaltered. As the autoregressive component has typically the highest weight in each VAR equation, any "indirect" effect induced by changing other coefficients may dominate the dynamic effects of the counterfactual experiment. Setting this indirect effect to zero ensures a clear interpretation of the results with respect to the objective of the experiment (see Canova and Gambetti, 2009). Similarly, the second experiment—the passive stance—involves the lagged coefficients on spending and taxes in the tax equation; the shocks to the spending coefficients are now plus one-standard deviation and the ones affecting the autoregressive component are again zero.

3. Bayesian Estimation

Starting with the papers of Cogley and Sargent (2005) and Primiceri (2005), structural time-varying parameter VARs with a recursive identification scheme have spread into the macroeconometric literature, especially with applications to monetary policy. Their method is appealing because it estimates the joint posterior of all parameters in the model, a distribution from which it is difficult to sample directly, by splitting the problem into smaller blocks. The parameters within each block can then be drawn from the conditional distributions through Gibbs sampling.

The Gibbs sampler is a variant of a Markov chain Monte Carlo (MCMC) algorithm: it exploits the principle that it is typically easier to sample from a lower dimensional distribution, conditional on other parameters (i.e. the blocks). Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin (1995, chap. 11) show that the stationary distribution of the Markov Chain generated by the Gibbs sampler is the joint distribution we are looking for. Furthermore, MCMC algorithms yield smoothed estimates of the time-varying parameters as they use information based on the entire set of observations. Compared to particle filters, smoothing methods lead to more efficient estimates when, like in this paper, the interest is in the evolution of the observable states (see, e.g., Christopher A. Sims, 2001; Primiceri, 2005).

As a notational convention, a generic vector or matrix x^τ consists of a sequence of observable variables or estimates up to time τ, i.e. $x^\tau = \{x_t\}_{t=1}^{\tau}$. I further express a realization z_t conditional on an information set, say, x^τ as $z_{t|\tau}$ and, likewise, I abbreviate the conditional mean and variance of an arbitrary parameter θ as $\theta_{t|\tau}$ and $V_{t|\tau}^{\theta}$. The function $p(\cdot)$ denotes a generic density and $\dim(\cdot)$ specifies the dimension of a vector.

3.1. Priors

An obvious choice to calibrate the priors are simple estimates from time-invariant ordinary least square regressions on (1) and (5). Such a strategy has already been used by Cogley and Sargent (2005) and Primiceri (2005), among others. In its original form this strategy requires to run these auxiliary regressions on a training sample that covers data which are then discarded for the main analysis. As quarterly observations for federal debts are only recorded after 1970 sacrificing, say, ten years of data for a training sample throws away a lot of information and might leave us with a too short sample. If a training sample is not available Canova and Matteo Ciccarelli (2009) suggest to estimate the features of the priors on the entire sample, i.e. on data from 1970:1 until 2010:3 (see also Kirchner, Cimadomo and Hauptmeier, 2010). As a side effect, using a "full sample" prior, minimizes the uncertainty involved in choosing proper priors. To denote the time-invariant estimator I will use "hats".

The details for the specification of the prior densities follow Canova and Gambetti (2009) and are quite similar to the ones in Primiceri (2005) and other papers. For the initial states of all time-varying parameters the priors $p(\beta_0)$, $p(\tilde{\alpha}_0)$ and $p(\log\sigma_0)$ are normally distributed, while the hyperparameters $p(Q)$, $p(S_i)$ and $p(W)$ have an inverse Wishart distribution. Given the laws of motion (12), (13) and (14), this choice of prior distributions for the initial states and the hyperparameters leads to normal priors for the entire sequences $\{\beta_t,\, \tilde{\alpha}_t,\, \log\sigma_t\}_{t=1}^{T}$. Like the normal distribution the Wishart distribution requires two input arguments, the scale factor and the degrees of freedom. For the prior to be proper the degrees of freedom must exceed the dimension of the respective hyperparameter at least by one; a choice of "just one" puts as little weight as possible on the prior. As the inverse Wishart distribution is a conjugate prior for the covariance matrix of the corresponding time-varying parameters β_t, $\tilde{\alpha}_t$ and $\log\sigma_t$ the scale factor has to be a multiple of the time-invariant covariance used to calibrate the prior for the initial states and the degrees of freedom. Bringing everything together we have the following set of prior densities for the parameters,

$$
\begin{aligned}
p(\beta_0) &= N\big(\hat{\beta},\, \mathrm{Var}(\hat{\beta})\big),\\
p(\tilde{\alpha}_0) &= N\big(\hat{\tilde{\alpha}},\, \mathrm{Var}(\hat{\tilde{\alpha}})\big),\\
p(\log\sigma_0) &= N\big(\log\hat{\sigma},\, I_k)\big),
\end{aligned}
\tag{21}
$$

and the hyperparameters

$$
\begin{aligned}
p(Q) &= IW\Big(\big(0.0003\times(\dim(\hat{\beta})+1)\times \mathrm{Var}(\hat{\beta})\big)^{-1},\, \dim(\hat{\beta})+1\Big),\\
p(S_i) &= IW\Big(\big(0.001\times(i+1)\times \hat{S}_i\big)^{-1},\, i+1\Big), \quad i=1,\dots,k-1,\\
p(W) &= IW\Big(\big(0.001\times(\dim(\hat{\sigma})+1)\times I_k\big)^{-1},\, \dim(\hat{\sigma})+1\Big),
\end{aligned}
\tag{22}
$$

in which the variance $\hat{S}_i$ refers to the i-th block of S in (15) and the variance for $\log\sigma_0$ and W is arbitrarily chosen to be the identity matrix. The factors 0.0003 and 0.001 and the degrees of freedom in the prior specification correspond to the values in Canova and Gambetti (2009) and transforms the initial informative prior choice $\mathrm{Var}(\hat{\beta})$ and $\hat{S}_i$ into a diffuse and uninformative one where more weight is on the sample information. This practice is more or less standard in the TVP-VAR literature, although slight differences can be found for $p(Q)$. Primiceri (2005), for instance, uses a factor of 0.0001 and $2 \times \dim(\beta)$ for the degrees of freedom. The results for these two prior specifications are, however, very similar; see Cogley and Sargent (2005) for another paper that specifies the prior for $p(Q)$ as in (22).

3.2. Sampling Algorithm

The Gibbs sampling algorithm set forth here specifies three blocks of conditional distributions for all parameters in the model: the coefficient states β_t and Q; the covariance states $\tilde{\alpha}_t$ and S; and the volatility states σ_t and W. The first two blocks can easily be cast into a linear and Gaussian state space form and therefore the standard algorithm for Gibbs sampling of Chris K. Carter and Robert Kohn (1994) can be used. Drawing volatility states is a bit more tricky as they have a nonlinear and nonnormal state space form. Sangjoon Kim, Neil Shephard and Siddhartha Chib (1998) provide a linear and approximately Gaussian reformulation of the problem with the advantage of restoring the assumptions needed for the standard sampling algorithm to work. The approximation is necessary because the linear transformation leads to innovations in the observation equation that are distributed as $\log \chi^2(1)$. Following Kim, Shephard and Chib (1998), I approximate this $\log \chi^2$ distribution with a mixture of seven normals. The indicator matrix s^T defines, out of the seven components, the selection of normal approximations for these innovations over $t = 1, \ldots, T$.

Step 1: Coefficient states $p\left(\beta^T | y^T, \tilde{\alpha}^T, \sigma^T, s^T, V\right)$ and algorithm in detail. — Equations (11) and (13), rewritten here for convenience,

$$y_t = X_t' \beta_t + u_t \quad \text{and} \quad \beta_t = \beta_{t-1} + \nu_t, \tag{23}$$

constitute a state space model in which both u_t and ν_t are normally distributed with a zero mean and variances Ω_t and Q. Further, the block diagonal structure of (15) assumes that u_t and ν_t are mutually uncorrelated. Now, conditional on the data, $\tilde{\alpha}^T$, σ^T and V the variance Ω_t in the observation equation is known from (9) and we can therefore generate the whole sequence

β^T as in Lemma 2.1 of Carter and Kohn (1994):

$$p\left(\beta^T|y^T,\tilde{\alpha}^T,\sigma^T,s^T,V\right)=p\left(\beta_T|y^T,\tilde{\alpha}^T,\sigma^T,s^T,V\right)\prod_{t=1}^{T-1}p\left(\beta_t|\beta_{t+1},y^t,\tilde{\alpha}^t,\sigma^T,s^T,V\right). \tag{24}$$

Then, to get β^T from $p\big(\beta^T|y^T,\cdots\big)$ we, first, generate β_T from $p\big(\beta_T|y^T,\cdots\big)=N\big(\beta_{T|T},V^\beta_{T|T}\big)$ and, second, for $t=T-1,\ldots,1$ we draw β_t from $p\big(\beta_t|\beta_{t+1},y^t,\cdots\big)=N\big(\beta_{t|t+1},V^\beta_{t|t+1}\big)$. Starting from $\beta_{0|0}=\hat{\beta}$ and $V^\beta_{0|0}=\mathrm{Var}(\hat{\beta})$ the Kalman filter recursion over $t=1,\ldots,T$, i.e.

$$\begin{aligned}
&\beta_{t|t-1}=\beta_{t-1|t-1},\\
&V^\beta_{t|t-1}=V^\beta_{t-1|t-1}+Q,\\
&\beta_{t|t}=\beta_{t|t-1}+V^\beta_{t|t-1}X_t\left(X_t'V^\beta_{t|t-1}X_t+\Omega_t\right)^{-1}\left(y_t-X_t'\beta_{t|t-1}\right)\quad\text{and}\\
&V^\beta_{t|t}=V^\beta_{t|t-1}-V^\beta_{t|t-1}X_t\left(X_t'V^\beta_{t|t-1}X_t+\Omega_t\right)^{-1}X_t'\beta_{t|t-1},
\end{aligned}\tag{25}$$

leads to a draw of β_T form the normal distribution using the elements $\beta_{T|T}$ and $V^\beta_{T|T}$ from the last recursion. We now plug the results of the filter and the draw of β_T into a reversed version of the Kalman filter to derive $\beta_{T-1|T}$ $V^\beta_{T-1|T}$. This backward updating delivers a draw for β_{T-1} and so forth until we arrive at β_1. Specially, the backward updating steps for $t=T-1,\ldots,1$ are

$$\begin{aligned}
&\beta_{t|t+1}=\beta_{t|t}+V^\beta_{t|t}\left(V^\beta_{t|t}+Q\right)^{-1}\left(\beta_{t+1}-\beta_{t|t}\right)\quad\text{and}\\
&V^\beta_{t|t+1}=V^\beta_{t|t}-V^\beta_{t|t}\left(V^\beta_{t|t}+Q\right)^{-1}V^\beta_{t|t}.
\end{aligned}\tag{26}$$

For more details on Gibbs sampling for state space models and the Kalman filter see Carter and Kohn (1994) and Brian D. O. Anderson and John B. Moore (1979).

So far nothing ensures that draws of β^T result in a stable VAR process. In fact, the use of data in level form with a more or less clear upward drift and nonstationary behavior leads hardly to any stable draw because a stable VAR process is by definition stationary (see Proposition 2.1 in Helmut Lütkepohl, 2005) As such, a strict "rule" as in Cogley and Sargent (2001) that discards every complete sequence of draws β^T where at least one draw β_t has an unstable VAR representation is simply infeasible. For the analysis of monetary policy, stability is a sensitive matter since the central bank's main objective is to maintain price stability. The advantage in a monetary policy VAR, however, is that the variables typically enter in first differences (most likely stationary form) and the rule has therefore less bite and will not slow down the sampling algorithm significantly. In a fiscal policy VAR, on the other hand, the whole stability issue may be less of a concern. As Kirchner, Cimadomo and Hauptmeier (2010) argue, certain episodes may be well described by fiscal instabilities. In every case, I take a compromise here and impose the stability rule on the growth rates of output, spending, taxes and federal debts. Specifically, I check the roots of the associated VECM polynomial of the VAR and discard every draw that has more than $k = 4$ roots in or on the unit circle.

Step 2: Covariance states $p\left(\tilde{\alpha}^T | y^T, \beta^T, \sigma^T, s^T, V\right)$. — Starting with the compact form of the structural model (5), we can derive the observation equation of the proper state space model from

$$A_t^* \left(y_t - X_t' \beta_t\right) = y_t^* = \tilde{A}_t e_t. \tag{27}$$

Conditional on β^T and the matrix of predetermined contemporaneous relations the adjusted residuals y_t^* are observable. As in Primiceri (2005), here is the point where the triangular form of the matrix $\tilde{A}_t$ with ones on the main diagonal can be conveniently used to rewrite (27) as observation equation

$$\begin{bmatrix} y_{g,t}^* \\ y_{t,t}^* \\ y_{x,t}^* \\ y_{d,t}^* \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ e_{g,t} & 0 & 0 & 0 & 0 & 0 \\ 0 & e_{g,t} & e_{t,t} & 0 & 0 & 0 \\ 0 & 0 & 0 & e_{g,t} & e_{t,t} & e_{x,t} \end{bmatrix} \begin{bmatrix} \tilde{\alpha}_{21,t} \\ \tilde{\alpha}_{31,t} \\ \tilde{\alpha}_{32,t} \\ \tilde{\alpha}_{41,t} \\ \tilde{\alpha}_{42,t} \\ \tilde{\alpha}_{43,t} \end{bmatrix} + \begin{bmatrix} e_{g,t} \\ e_{t,t} \\ e_{x,t} \\ e_{d,t} \end{bmatrix}. \tag{28}$$

and (12) serves as state equation for $\tilde{\alpha}_t$. Now, the block-diagonality of the covariance matrix S of the innovations ζ_t and block-triangular structure of the 4×6 matrix in (28), enables us to use the algorithm of Carter and Kohn (1994), explained in Step 1, in an equation-by-equation fashion. Specifically, given β^T and the triangular form of $\tilde{A}_t$, $e_{g,t}$ is predetermined and thus $\tilde{\alpha}_{21,t}$ can be drawn in the first equation. For the second equation we use the draw of $\tilde{\alpha}_{21,t}$ and predetermine $e_{t,t}$ such that we can obtain draws for the block $[\tilde{\alpha}_{31,t}, \tilde{\alpha}_{32,t}]$. Continuing this procedure of predetermining one structural shock at the time leads to draws for the block $[\tilde{\alpha}_{41,t}, \tilde{\alpha}_{42,t}, \tilde{\alpha}_{43,t}]$ in the third equation and so forth. The triangular structure of predetermined variables in the individual equations and the independence across the blocks of S restores the necessary assumption of a linear state space model in the Carter and Kohn (1994) algorithm.

TABLE 1—SELECTION OF THE MIXING DISTRIBUTION TO BE $\log\chi^2(1)$

j	p_j	m_j	v_j^2	j	p_j	m_j	v_j^2
1	0.00730	−10.12999	5.79596	5	0.34001	0.61942	0.64009
2	0.10556	−3.97281	2.61369	6	0.24566	1.79518	0.34023
3	0.00002	−8.56685	5.17950	7	0.25750	−1.08819	1.26261
4	0.04395	2.77786	0.16735				

Notes: Replication of Table 4 in Kim, Shephard and Chib (1998).

Step 3: Volatility states $p\left(\sigma^T|y^T,\beta^T,\tilde{\alpha}^T,s^T,V\right)$. — Drawing σ^T relies on the algorithm of Kim, Shephard and Chib (1998), a procedure to transform an otherwise nonlinear and nonnormal state space model into a linear and approximately normal one; the standard algorithm of Carter and Kohn (1994), as laid out in step 1, is the again available. The observation equation can be written as

$$\tilde{A}_t^{-1}A_t^*\left(y_t - X_t'\beta_t\right) = \Sigma_t\varepsilon_t. \tag{29}$$

Given $y^T,\beta^T,\tilde{\alpha}^T$ the right-hand side is observable and is nothing else than the set of identified structural shocks, e_t, of Step 2. Since I have defined the law of motion (14) for the diagonal entries of Σ_t as a geometric random walk, we can convert (29) into the appropriate form by squaring and taking the logarithm. We obtain the linear state space model

$$e_t^* = 2\log\sigma_t + \xi_t \quad \text{and} \quad \log\sigma_t = \log\sigma_{t-1} + \eta_t, \tag{30}$$

in which $e_{i,t}^* = \log\left(e_{i,t}^2 + 0.001\right)$ and $\xi_{i,t} = \log\left(\epsilon_{i,t}^2\right)$ for $i = (g,t,x,d)$; the offset constant 0.001 deals with very small values of $e_{i,t}^2$ as in Kim, Shephard and Chib (1998); and the innovation ξ_t follows a $\log\chi^2(1)$ distribution. While this conversion restores the linearity assumption the distributional form of ξ_t still precludes direct and simple inference. Kim, Shephard and Chib (1998) show how to accurately approximate the $\log\chi^2(1)$ distribution through a matched mixture of normal distributions,

$$f\left(\xi_{i,t}\right) \approx \sum_{j=1}^{7} p_j N\left(\xi_{i,t}|m_j - 1.2704,\, v_j^2\right), \qquad i = (g,t,x,d), \tag{31}$$

in which $N\big(\xi_{i,t}|m_j-1.2704,\, v_j^2\big)$ denotes the density function of a normal distribution with mean $m_j-1.2704$ and variance v_j^2. Values for p_j, m_j and v_j^2 are reproduced in Table 1. Conditional on s^T we can draw a value for $\xi_{i,t}|s_{i,t}=j \sim N\big(m_j-1.2704,\, v_j^2\big)$ and proceed as in Step 1 to draw $\log\sigma_{i,t}$ for all i and t. Given these draws of $\xi_{i,t}$ we independently sample each $s_{i,t}$ from the discrete density $\Pr\left(s_{i,t}=j|e_t^*,\log\sigma_{i,t}\right)$, a density which is proportionally determined from the normal density $N\big(e_t^*|2\log\sigma_{i,t}+m_j-1.2704, v_j^2\big)$.

Step 4: Hyperparameters. — The inverse Wishart is a convenient choice for the prior distribution of the innovation variances V in (15), i.e. the hyperparameters Q, W and the blocks of S. Since the parameters β^T, $\tilde{\alpha}^T$ and σ_T are mutually uncorrelated draws from a normal distribution the posterior distribution of each hyperparameter is also inverse Wishart. Conditional on β^T, $\tilde{\alpha}^T$, σ^T, s^T and y^T the innovations in (15) become observable and it is therefore relatively easy to draw the hyperparameters from the inverse Wishart. The scale matrix and the degrees of freedom, the two factors that fully specify the inverse Wishart, are based on the choice for the prior distribution and take the form

$$
\begin{gathered}
\Big(0.0003\times(\dim(\hat{\beta})+1)\times\mathrm{Var}(\hat{\beta})+\sum_{t=1}^{T}\Delta\beta_t\Delta\beta_t'\Big)^{-1} \quad\text{and}\quad \dim(\hat{\beta})+1+T,\\
\Big(0.001\times(i+1)+T\times\hat{S}_i\Big)^{-1} \quad\text{and}\quad i+1+T, \quad i=1\ldots,k-1,\\
\Big(0.001\times(\dim(\hat{\sigma})+1)\times I_k+\sum_{t=1}^{T}\Delta\sigma_t\Delta\sigma_t'\Big)^{-1} \quad\text{and}\quad \dim(\hat{\sigma})+1+T.
\end{gathered}
\tag{32}
$$

in which $\hat{S}_i$ denotes the variance of the i-th block of S in (15).

For the counterfactual analysis, Step 1 needs to be slightly modified. Everything else being as just laid out, the shocked and shock-free realizations of the impulse response function (20) come from draws of β_{t-1}. The sequence of parameters β_{t+h}, $h=0,1,\ldots,H$, follows then from (13) and (19) either with $\tilde{\omega}_t=\delta$ or $\tilde{\omega}_t=0$.

The Gibbs sampling algorithm is now complete. Iterations on Steps 1 to 4 produce a set of draws from the conditional distributions that converge in the limit to the joint posterior distribution of all the parameters in the model (see, e.g., Alan E. Gelfand and Adrian F. M. Smith, 1990). I perform $100,000$ iterations from which I discard the first $50,000$ and save only every fifth draw of the remaining $50,000$ draws. This "thinning" practice breaks the autocorrelation of the draws since draws from a Markov chain are typically not independent.

3.3. Convergence Diagnostics of the Markov Chain

From theoretical work such as Gelfand and Smith (1990) we know that the Gibbs sampler converges to the "true" joint posterior distribution as the number of iterations go to infinity. Whether this property holds in the underlying problem with a finite number is an important question which I address here. Intuitively, convergence of the Markov chain slows down the more complicated the conditional distribution gets.

I implement three MCMC convergence diagnostics for the $10,000$ saved draws of each parameter and hyperparameter: the sample autocorrelation; the measure of John Geweke (1992); and the Adrian E. Raftery and Steven M. Lewis (1992) diagnostic. Table 2 reports the results of the diagnostic checks for each of the $8,732$ parameters in the model. Because of the sheer amount of parameters the table shows summary statistics, grouped into hyperparameters V, coefficients β^T, covariances $\tilde{\alpha}^T$ and volatilities σ^T. Moreover, each summary statistic reports two values based on the first and last $1,500$ draws from the $10,000$ saved iterations. This testing strategy adds another layer to the formal MCMC diagnostics: if the Markov chain is in an equilibrium state the means of these two splits should be roughly equal.

The 20-th-order sample autocorrelation summarized in Panel A of Table 2 show a relatively low degree of autocorrelation. Only a few hyperparameters V exhibit statistics higher than 0.2. The draws are therefore almost independent, an indication for the efficiency of the algorithm and for accurate posterior estimates. Related to that is the inefficiency factor, as measured by the inverse of the relative numerical efficiency statistic of Geweke (1992) with a 4% tapered window for the estimation of the spectral density at frequency zero. If the draws come from an independent and identically distribute (iid) sample, drawn directly form the posterior distribution, the inefficiency factor has a value of one. For instance, in Panel B of Table 2 the mean value of 10.94 for the last $1,500$ draws of the hyperparameters V indicate that about eleven times as many draws are necessary to achieve the same numerical efficiency of an iid set of draws. Since only values above 20 are considered to be critical and 10.94 is the largest one here, Table 2 confirms the iid nature of the draws. Finally, Raftery and Lewis (1992) provide a measure of the number of draws actually required to achieve a certain accuracy of the posterior summaries of the Markov chain. I set the parameters for this test such that a nominal reporting based on a 95% interval using the 0.025 and 0.975 quantile points leads to an accuracy of the posterior values of 0.025 to the left and right of the specified quantiles in the cumulative distribution function. The probability of attaining this accuracy is 95%. The maximum number over the whole parameter space for the Raftery and Lewis (1992) diagnostic is 4,818 and thus well below the 10,000 draws used in the analysis. All three MCMC convergence diagnostics do not indicate any problems with the Gibbs sampler.

Table 2—Convergence Diagnostics of the Markov Chain

	Median		Mean		Min		Max		10-th		90-th	
A. 20-th-Order Sample Autocorrelations												
V	−0.02	0.04	−0.02	0.04	−0.31	−0.05	0.44	0.49	−0.15	−0.01	0.13	0.08
β^T	0.00	0.00	0.00	0.00	−0.20	−0.03	0.18	0.05	−0.07	−0.01	0.07	0.02
$\tilde{\alpha}^T$	−0.04	0.01	−0.03	0.01	−0.14	−0.02	0.10	0.03	−0.11	−0.01	0.07	0.02
σ^T	−0.02	0.00	−0.02	0.01	−0.17	−0.03	0.11	0.09	−0.07	−0.01	0.05	0.02
B. Inefficiency Factor												
V	5.74	10.67	5.74	10.94	3.79	4.16	9.54	41.34	4.54	7.36	6.86	14.50
β^T	1.11	1.43	1.15	1.50	0.50	0.61	2.49	4.35	0.86	1.04	1.50	2.02
$\tilde{\alpha}^T$	0.94	1.08	0.96	1.30	0.70	0.77	1.36	2.69	0.85	0.83	1.13	2.21
σ^T	1.31	1.66	1.38	1.87	0.74	0.78	3.57	10.28	0.98	1.10	1.87	2.64

Notes: Summary of the distributions of the 20-th-order sample autocorrelations and the inefficiency factors (the inverse of Geweke's (1992) measure of relative numerical efficiency with a 4% tapering of the spectral window at frequency zero) for the whole parameter space. "10-th" and "90-th' denote the 10-th and 90-th percentiles. Each statistic has two entries which refer to statistics based on the first and last 1,500 draws out of the saved 10,000. The discarded burn-in draws are 50,000 and the thinning factor is five.

4. Results

I center the discussion of the results, especially the observed changes over the last 40 years, around three topics: the volatility of government spending and tax shocks, the propagation of these shocks and counterfactual fiscal policy scenarios.

4.1. Volatility of the Fiscal Shocks

Figure 2 shows the time profile of the median and the interval containing 68% of the posterior distribution of the standard deviation of the estimated fiscal policy shocks. For the government spending shock the median and the 68% interval are fairly stable over time while for tax shocks these statistics are, on average, lower in the 1980s and 1990s than at the beginning and the end of the sample. The picture of a relatively high volatility in the 1970s, with a peak around 1975, and in the 2000s is consistent with Romer and Romer's (2010) narrative analysis of tax shocks. The 1970s and early 1980s were periods of frequent and large tax changes, such as presidents Ford and Reagan's tax cuts, mainly aimed to boost long-run growth or to counteract economic conditions. Until the Bush tax cuts of the early 2000s and the tax measures included in the 2008-2009 stimulus packages there were only a few and relatively modest deficit-driven tax changes.

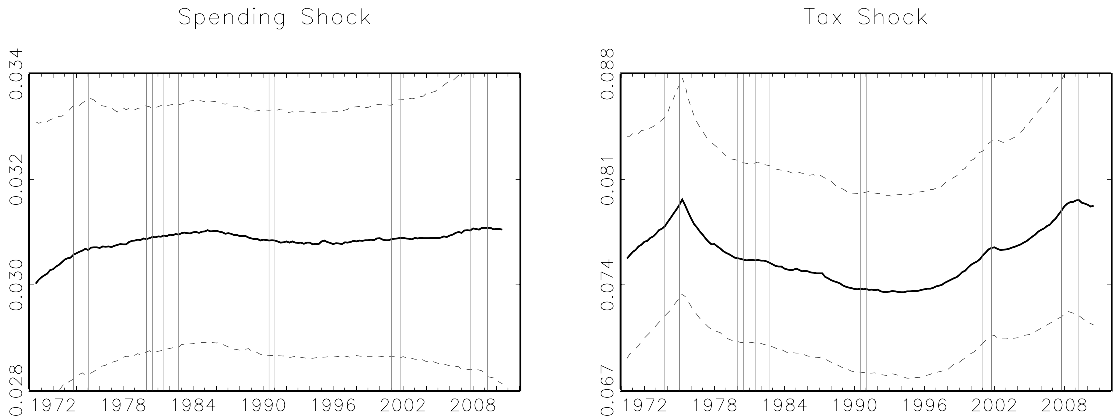

Figure 2. Standard Deviation of the Structural Shocks

Notes: The two graphs displays the time-varying volatility parameters $\sigma_{1,t}$ and $\sigma_{2,t}$ in (8). Superimposed are NBER-dated U.S. recession episodes.

4.2. The Dynamic Responses to Fiscal Shocks

Figures 3 and 4 display the systematic responses to fiscal shocks. I focus on three specific periods: 1975:2, 1991:2 and 2009:1. Although all of them represent NBER-dated troughs of U.S. recessions the sole objective is to uncover changes in the transmission of fiscal policy shocks over the last 40 years. Since I use smoothed estimates based on the entire set of observations the responses in, say, 1973:4 and 1975:2 are hardly to distinguish from each other. Overall, the impulse responses support, perhaps with a few exceptions, the common belief about the changing transmission mechanism of fiscal policy. Robert M. Solow (2005) brings this change to the point: "[t]he use of fiscal policy as a stabilization device has all but vanished [...] in the United States."

The output response to a spending shock (Figure 3) was more effective in (and around) 1975:2 when there was considerable slack in the U.S. economy; the general U-shaped pattern is in line with the results of Blanchard and Perotti (2002). While the spending response is relatively persistent and similar over time, tax revenues fall below zero after an initial increase and slowly revert to trend. This U-shape is mainly driven by the automatic adjustment of tax collections to changes in output but the pattern is somewhat different over time, reflecting changes in the Taylor-type rule describing tax-policy making (i.e. the structural tax equation). Around 1975:2, a change in spending today leads to a faster decline of tax revenues tomorrow. This faster decline comes, however, at the cost of higher future debts.

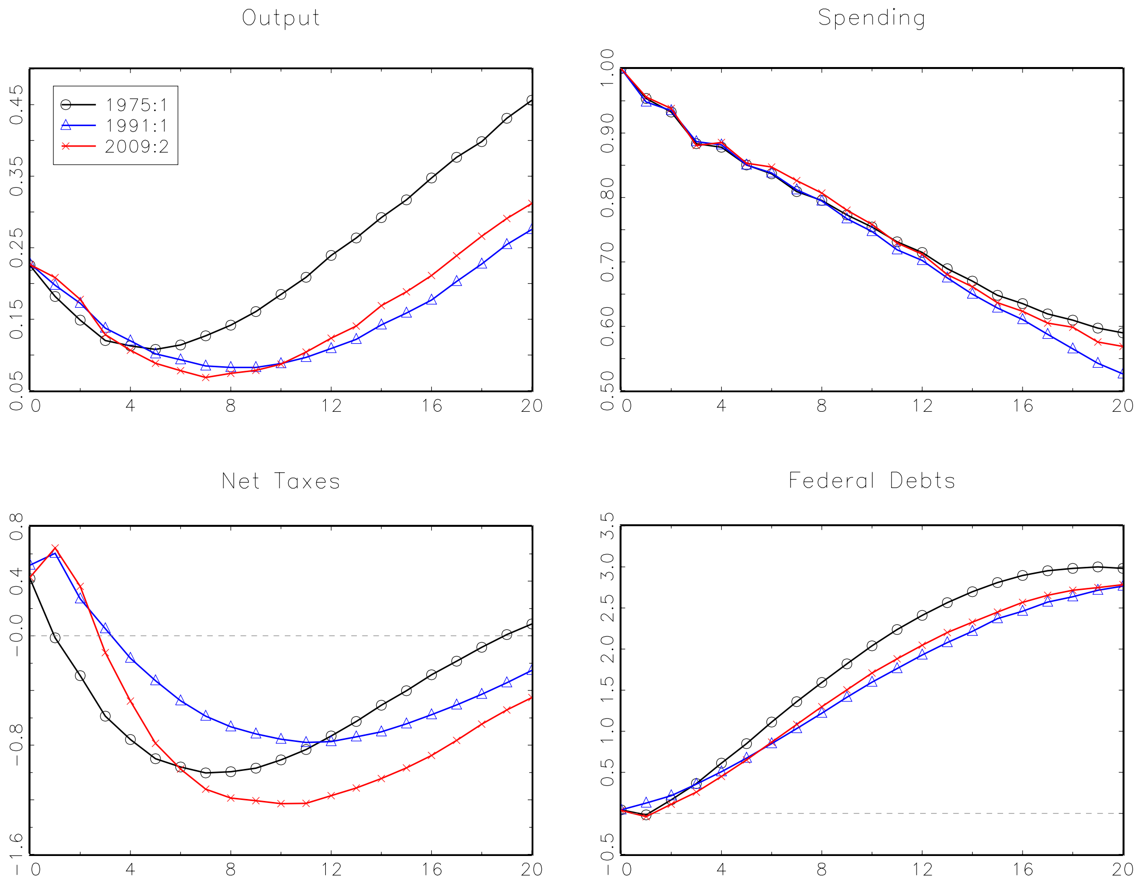

Figure 3. Responses to Spending Shocks

Notes: Impulse responses of output, government spending, net taxes and federal debts to spending shocks in 1975:1, 1991:1 and 2009:2. Responses expressed in percentage changes.

In Figure 4, the output effects to tax shocks are, perhaps, the exception to the mentioned conventional wisdom of a declining fiscal policy effectiveness. A tax shock leads to a higher (negative) output response in 1991:1 than in 1975:1. The years around 1991 was a period of several deficit driven tax changes aimed to bring the ever rising debt-to-output ratio to a halt (see right panel in Figure 1). More persistent tax revenues and lower spending, although the spending effect is in general quite small, reduce debt levels more effectively than in other periods. This debt reduction motive of the government has, of course, detrimental effects on output.

Much of the debate about fiscal policy effectiveness centers around the size of spending and tax multipliers. Figure 5 displays the evolution of the output responses at selected horizons

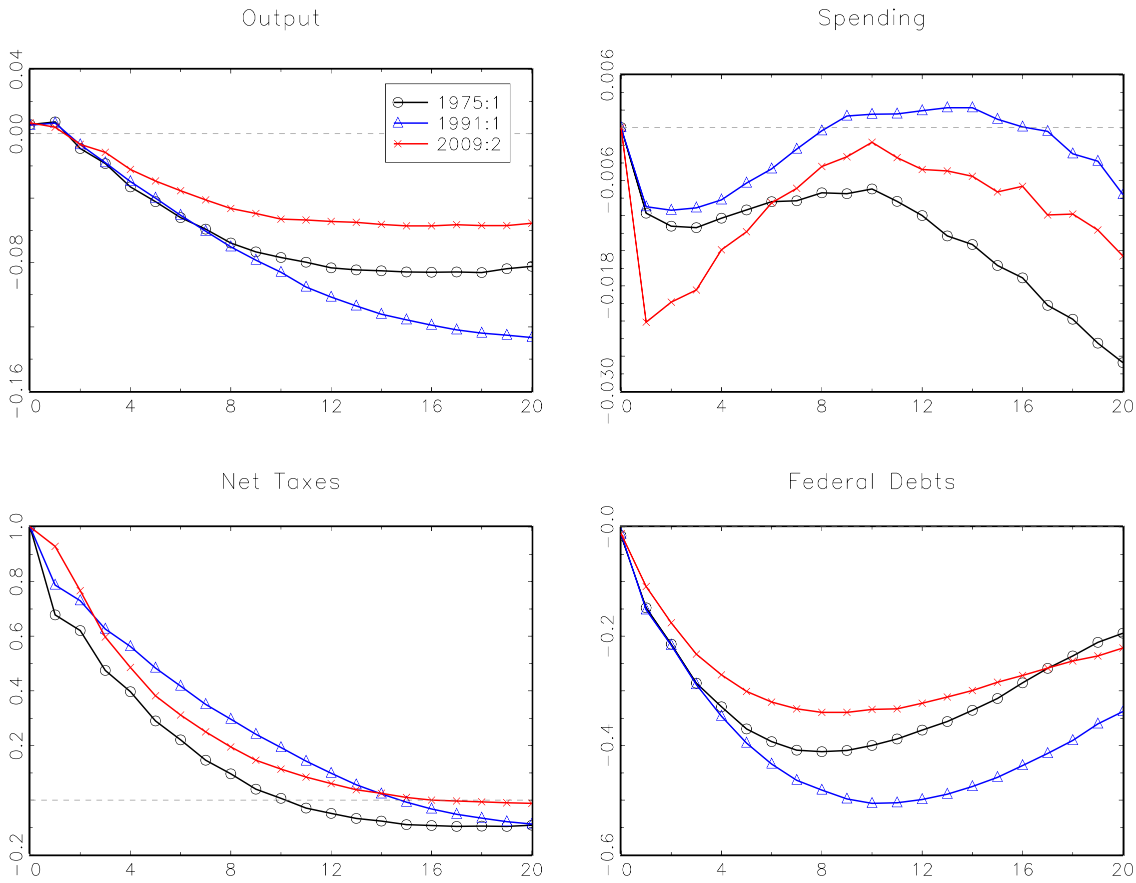

FIGURE 4. RESPONSES TO TAX SHOCKS

Notes: Impulse responses of output, government spending, net taxes and federal debts to tax shocks in 1975:1, 1991:1 and 2009:2. Responses expressed in percentage changes.

from 1970:3 until 2010:3 and expressed as dollar for dollar changes. To convert the original elasticity estimates from before into derivatives, I divide the elasticities by the spending-to- and tax-to-output ratio prevailing at time t. These are the ratios plotted in the left panel of Figure 1. For better comparison of the two multiplier effects, I compute the tax multiplier based on negative shocks.

Three results stand out from the multiplier analysis. First, the spending multiplier in the 1970s is, on average, only more effective in the longer run: the government buys with one dollar about 1.60 dollars of output four years out. Second, the impact spending multiplier increases almost steadily over time, from one dollar in 1970 to 1.25 dollars in 1998. This effect arises, however, mostly through the declining spending share (see Figure 1). In every case, the size of the spending multipliers is reasonable throughout taking values between 0.50 and 1.60 dollars before 1980 and between 0.50 and 1.25 dollars after 1980. Table 2 in Robert E. Hall (2009) provides a summary of several time-invariant VAR estimates: there the multipliers range from 0.50 up to 1.20 dollars. Finally, tax multipliers lie consistently below the spending multipliers. Starting with a value of 0.50 dollars in 1970 the effect after four years reaches its peak at roughly 0.80 dollars in 1983 and decreases thereafter until it reaches again the 0.50 dollar mark in 2010. For the other horizons the effects lie between zero and 0.40 dollars with no obvious trends. The comparison of tax multipliers with other papers is flawed because recent papers such as Romer and Romer (2010) and Perotti (2011) use a narrative approach to identify tax shocks and get much larger effects. For that reason, I plan to incorporate narrative measures of tax shocks in a fiscal TVP-VAR in future work.

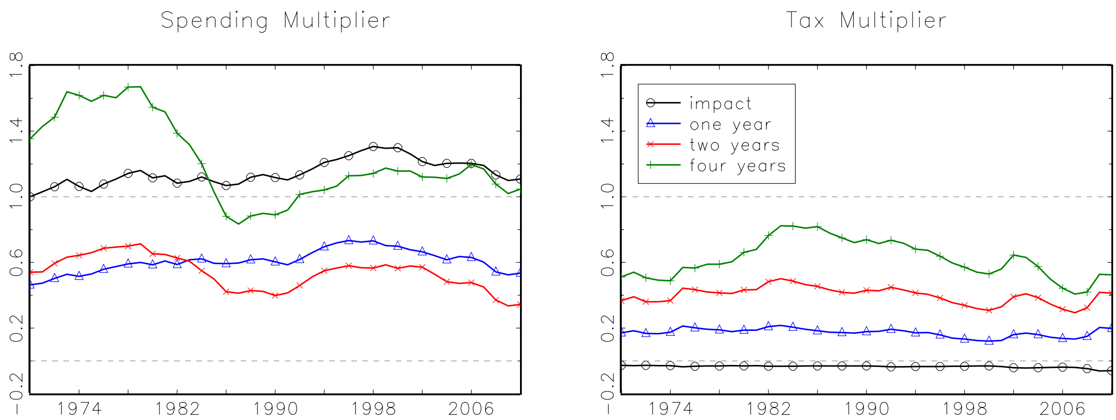

FIGURE 5. MULTIPLIERS

Notes: The spending and tax multipliers display the output responses expressed as dollar for dollar changes. Conversion from percentage changes into dollar changes based on the spending-to- and tax-to-output ratio prevailing at time t (see ratios in Figure 1). For better comparison with the spending multiplier, I use a negative tax shock to compute the tax multiplier.

4.3. Fiscal Policy Counterfactuals

The objective of the counterfactual experiments laid out in Section 2.3 is to simulate a government that reduces the level of debts in two different ways, an active and passive one. In the first experiment I let the government cut spending more aggressively whenever we observe rising debts and, in the second one, the government adjusts taxes more swiftly in response to higher expenditures. Scenario one requires the government to directly control the level of debts through tighter spending constraints and has a stronger incentive to increase its own efficiency. Under scenario two, on the other hand, the government just levies enough taxes in order to pay for whatever expenditures they made in the previous periods. The questions is then how the different incentives the government has in the active and passive stance will spread over the private sector and how, as a result, it will affect output.

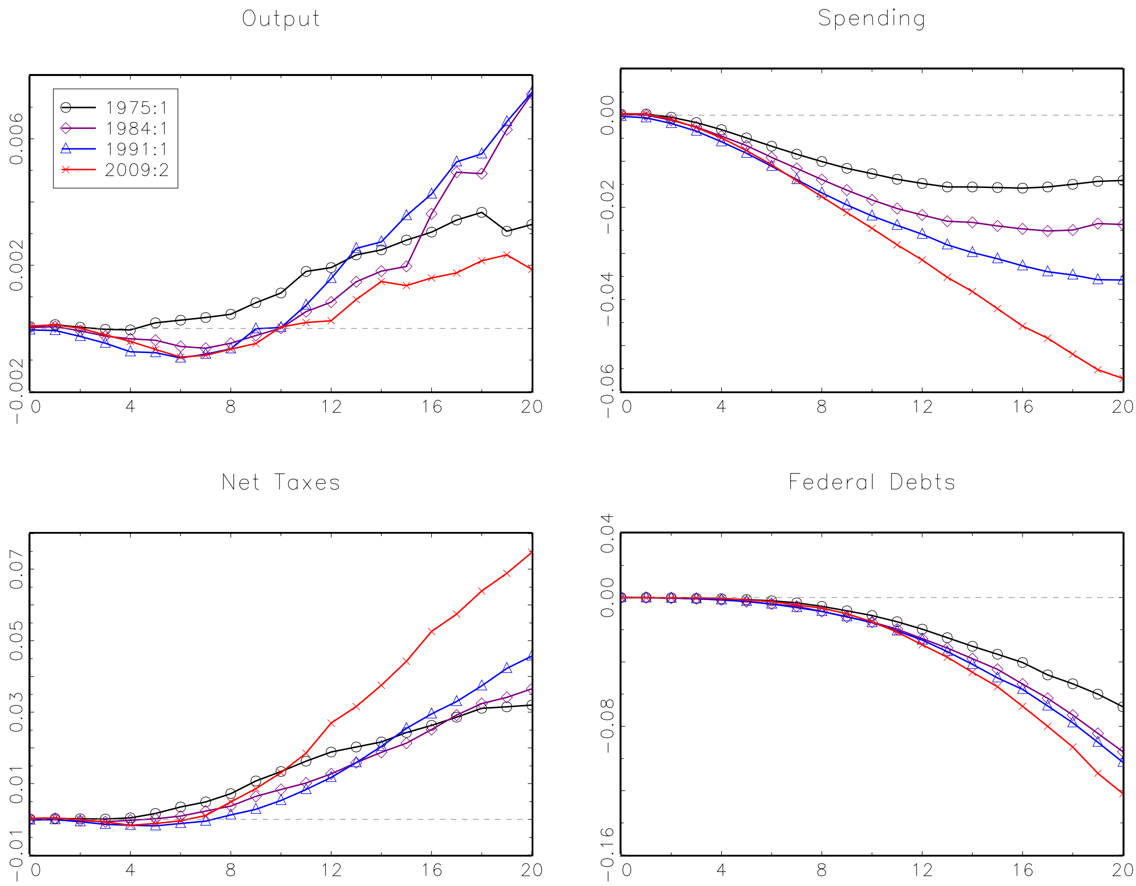

FIGURE 6. ACTIVE GOVERNMENT STANCE

Notes: Counterfactual responses of output, government spending, net taxes and federal debts in 1975:1, 1984:1, 1991:1 and 2009:2. The experiment simulates a government that reduces debts by means of more aggressive spending cuts. Computed along the lines of Section 2.3 and Canova and Gambetti (2009). Responses expressed in percentage changes.

The design of the counterfactual follows Canova and Gambetti (2009) and circumvents the Lucas critique by accounting for the effect a change in one coefficient has on the whole coefficient structure β_t through the estimated correlation Q among the coefficients. I display the counterfactual responses in Figures 6 and 7 at four specific dates: the three recession trough dates used above, 1975:1, 1991:1, 2009:2 and in addition 1984:1. The impulse response function in (20) implies a dependence on the local history of the variables. I add this specific date because it represents a time when all four variables were growing rapidly. The additional date, therefore, provides a natural counterpart to the three trough dates.

The responses in Figures 6 and 7 confirm the hunch that the two government stances imply different effects on output. In the active stance the spending cuts have no adverse output effects in the longer run: it positively affects the private sector and outweighs the stresses and strains from the debt reduction. When the government is in the passive stance, output decreases: even though spending increases the government raises taxes swiftly and puts any positive incentives for the private sector on hold.

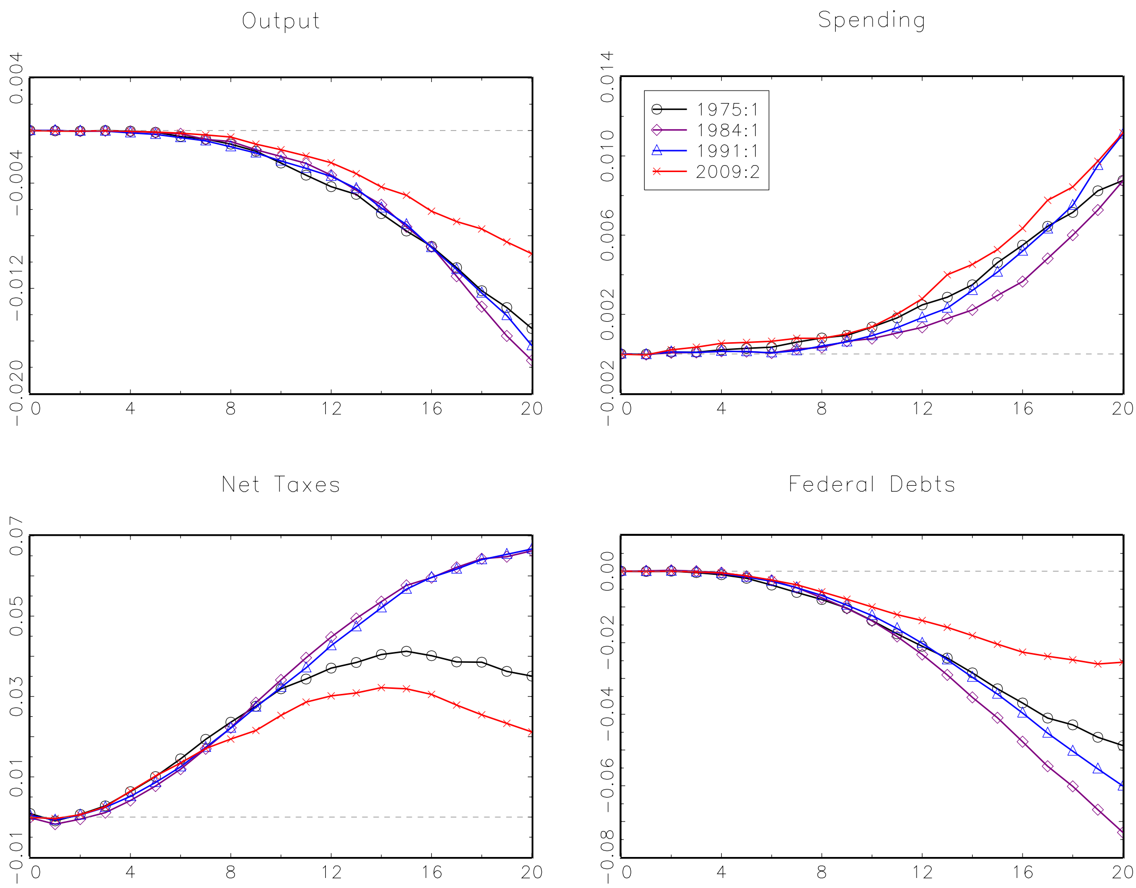

Figure 7. Passive Government Stance

Notes: Counterfactual responses of output, government spending, net taxes and federal debts in 1975:1, 1984:1, 1991:1 and 2009:2. The experiment simulates a government that reduces debts by adjusting taxes more quickly in response to higher expenditures. Computed along the lines of Section 2.3 and Canova and Gambetti (2009). Responses expressed in percentage changes.

The counterfactual responses in the recent 2009:2 period differ somewhat from other periods. In Figure 6 the government reduces spending and, at same time, increases taxes by more than in 1975:1, 1984:1 and 1991:1. This result is emblematic for the severe recession in 2008 and 2009 with extreme changes in all four variables which, in turn, require larger counterfactual responses of the public sector variables in order to achieve a certain goal. Only output lags behind: the counterfactual output response lies below the others at horizons beyond ten quarters. Also in the passive stance in Figure 7 one sees a much smaller negative output response in the 2009:2 period. The spending increase is accompanied by a less pronounced rise in taxes and a smaller reduction of debts, thus the subdued effects on output.

The general message here is clear. A government that actively reduces debts by cutting spending in a credible way provides enough incentives for the private sector to pitch in for the government in the longer run. Passively reducing debts by just controlling the deficit through tax adjustments does not seem to be good policy-making.

5. Conclusions

In this paper, I contribute to the recent literature on the effects of fiscal policy by highlighting the time-variation in the transmission of government spending and tax shocks. My analysis is cast in the time-varying parameter structural VAR framework of Primiceri (2005) which, with a few exceptions, has been mainly used so far to study monetary policy. Specifically, I model fiscal policy and the private sector behavior of the U.S. economy over the last four decades, including data on output, government spending, net taxes and federal debts. Estimation relies on an efficient Markov chain Monte Carlo algorithm, Gibbs sampling in particular, for the numerical evaluation of the posterior distributions.

The main results accord well with the conventional wisdom of declining effectiveness of changes in fiscal policy, with one qualification for taxes. Unlike other periods, the late 1980s and early 1990s were characteristic for deficit-driven tax changes with the objective to reverse the course of the surging debt-to-output ratio. The TVP-VAR uncovers this mid-period as the one when tax policy was most effective, especially with respect to reducing the level of debts. Overall, the amount and pattern of time-variation observed, makes a TVP-VAR an attractive and natural choice for the empirical characterization of changes in fiscal policy.

From a methodological point of view the paper innovates upon the TVP-VAR literature in one important aspect. It implements the widely used Blanchard and Perotti (2002) method to identifying fiscal policy shocks into a TVP-VAR. The twist is, simply put, that identification requires a non-recursive structure of the contemporaneous impact matrix, whereas Primiceri's (2005) framework relies on a triangular shape of that matrix. While Pereira and Lopes (2010) are the first who provide a solution to this aspect, my reformulation of the problem is more compatible with the estimation algorithm of Primiceri (2005). Whether the method of Blanchard and Perotti (2002) is the best way to identifying the "true" underlying government spending and tax shocks was beyond the scope of this paper. In every case, their is plenty of room for more ingenuity. An immediate extension would be to identify the shocks through sign restrictions as in Mountford and Uhlig (2009), although eliciting better information on impulse responses is by no means guaranteed as the identifying restrictions provide very weak information. On the other hand, combing the new narrative-based tax measures, most notably Romer and Romer (2010), with TVP-VARs is, perhaps, a more promising avenue for future research in order to deal with identification issues. Especially the ones arising through fiscal foresight.

In addition to the methodological contribution, I use the counterfactual policy design of Canova and Gambetti (2009) to study the effects of two different ways to reduce the level of debts: actively by cutting spending and passively through budget surpluses obtained from tax adjustments in response to past expenditure levels. In that respect, this counterfactual analysis also bears on the policy debate about what should follow after the recent, mostly deficit-financed, stimulus packages. As one may perhaps expect, the active government policy stance has hardly any adverse effects on output. In fact, output tends to increase in the longer run. The passive stance, on the other hand, has no positive effects on private sector behavior and, consequently, output decreases. The differences arise through the way the government actions affect private incentives. A government that reduces the debt burden in a credible and active way and, in the same time, may increase its own efficiency provides enough positive incentives for the private sector to more than compensate for the reduced public expenditures.

REFERENCES

Anderson, Brian D. O., and John B. Moore. 1979. *Optimal Filtering.* Englewood Cliffs, NJ: Prentice Hall.

Auerbach, Alan, and Yuriy Gorodnichenko. 2010. "Measuring the Output Responses to Fiscal Policy." NBER Working Paper No. 16311.

Benati, Luca, and Paolo Surico. 2008. "Evolving U.S. Monetary Policy and the Decline of Inflation Predictatbility." *Journal of the European Economic Association*, 6(2-3): 634–646.

Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics*, 117(4): 1329–1368.

Canova, Fabio, and Luca Gambetti. 2009. "Structural Changes in the US Economy: Is there a Role for Monetary Policy?" *Journal of Economics Dynamics and Control*, 33(2): 477–490.

Canova, Fabio, and Matteo Ciccarelli. 2009. "Estimating Multicountry VAR Models." *International Economic Review*, 50(3): 929–959.

Carter, Chris K., and Robert Kohn. 1994. "On Gibbs Sampling for State Space Models." *Biometrika*, 81(3): 541–553.

Cogley, Timothy, and Thomas J. Sargent. 2001. "Evolving Post-World War II U.S. Inflation Dynamics." In *NBER Macroeconomics Annual*, 331–373. Cambridge, MA: MIT Press.

Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and Volatilities: Monetary Policies and Outcomes in the post WWII US." *Review of Economic Dynamics*, 8(2): 262–302.

Favero, Carlo, and Francesco Giavazzi. 2007. "Debt and the Effects of Fiscal Policy." NBER Working Paper No. 12822.

Gelfand, Alan E., and Adrian F. M. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association*, 85(410): 398–409.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 1995. *Bayesian Data Analysis.* London: Chapman and Hall.

Geweke, John. 1992. "Evaluating the Accruacy of Sampling-Based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics*, ed. J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith, 169–193. Oxford: Oxford University Press.

Hall, Robert E. 2009. "By How Much Does GDP Rise If the Government Buys More Output?" *Brookings Papers on Economic Activity*, 2: 183–231.

Kim, Sangjoon, Neil Shephard, and Siddhartha Chib. 1998. "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models." *Review of Economic Studies*, 65(3): 361–393.

Kirchner, Markus, Jacopo Cimadomo, and Sebastian Hauptmeier. 2010. "Transmission of Government Spending Shocks in the Euro Area: Time Variation and Driving Forces." ECB Working Paper No. 1219.

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. "Impulse Response Analysis in Nonlinear Multivariate Systems." *Journal of Econometrics*, 74: 119–147.

Leeper, Eric M., and Tao Zha. 2003. "Modest Policy Interventions." *Journal of Monetary Economics*, 50(8): 1673–1700.

Leeper, Eric M., Todd B. Walker, and Shu-Chun Susan Yang. 2008. "Fiscal Foresight: Analytics and Econometrics." CAEPR Working Paper 2008-013.

Lütkepohl, Helmut. 2005. *New Introduction to Multiple Time Series.* Heidelberg: Springer.

Mountford, Andrew, and Harald Uhlig. 2009. "What are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics*, 24(6): 960–992.

Pereira, Manuel Coutinho, and Artur Silva Lopes. 2010. "Time Varying Fiscal Policy in the U.S." CEMAPRE Working Paper No. 1004.

Perotti, Robert. 2011. "The Effects of Tax Shocks on Output: Not so Large, But Not Small Either." NBER Working Paper No. 16786.

Perotti, Roberto. 2007. "In Search of the Transmission Mechanism of Fiscal Policy." NBER Working Paper No. 13143.

Primiceri, Giorgio E. 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies*, 72(3): 821–852.

Raftery, Adrian E., and Steven M. Lewis. 1992. "How Many Iterations in the Gibbs Sampler?" In *Bayesian Statistics*, ed. J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith, 763–773. Oxford: Oxford University Press.

Romer, Christina D., and David H. Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review*, 100(3): 763–801.

Sims, Christopher A. 2001. "Comment on Sargent and Cogley's: 'Evolving Post World War II U.S. Inflation Dynamics'." In *NBER Macroeconomics Annual*, 373–379. Cambridge, MA: MIT Press.

Solow, Robert M. 2005. "Rethinking Fiscal Policy." *Oxford Review of Economic Policy*, 21(4): 509–514.